

AD-A103 867

WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER

F/G 12/1

EXPLICIT FILTERS FOR DIFFUSIONS WITH CERTAIN NONLINEAR DRIFTS.(U)

AUG 81 D L O'CONNOR, J S BARAS, S I MARCUS

DAAG29-80-C-0041

UNCLASSIFIED

MRC-TSR-2254

NL

1-1
A-1
B-1



END
DATE
FILMED
0-81
DTIC

LEVEL

(2) 135

AD A103867

MRC Technical Summary Report #2254

EXPLICIT FILTERS FOR DIFFUSIONS WITH
CERTAIN NONLINEAR DRIFTS

D. L. Ocone, J. S. Baras
and S. I. Marcus

Mathematics Research Center
University of Wisconsin-Madison
610 Walnut Street
Madison, Wisconsin 53706

August 1981

Received April 30, 1981

DTIC
ELECTE
SEP 8 1981
S D
A

THIS FILE COPY

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

81 9 08 105

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

EXPLICIT FILTERS FOR DIFFUSIONS WITH CERTAIN NONLINEAR DRIFTS

D. L. Ocone¹, J. S. Baras² and S. I. Marcus³

Technical Summary Report #2254
August 1981

ABSTRACT

Let $x(t)$ be a diffusion satisfying the stochastic differential equation $dx(t) = f(x(t))dt + db(t)$, where $f'(x) + f^2(x) = ax^2 + bx + c$, $a > 0$. V. Beneš gave an explicit formula for the conditional density of $x(t)$ given $y(s)$, $0 \leq s \leq t$, where $y(s) = \int_0^s x(s)ds + w(t)$, when $w(\cdot)$ is a Brownian process independent of $x(\cdot)$. This result is extended and then applied to derive recursive filtering equations for estimating conditional moments $E\{x^n(t)|y(s), 0 \leq s \leq t\}$, for estimating polynomial functionals of $x(\cdot)$, and for smoothing.

AMS(MOS) Subject Classification No. 60G35, 62M20, 93E11

Key word: Nonlinear filtering

Work Unit No. 4 - Probability and Statistics

¹Mathematics Department and the Mathematics Research Center, University of Wisconsin-Madison, Madison, WI 53706.

²Electrical Engineering Department, University of Maryland, College Park, MD 20742.

³Department of Electrical Engineering, University of Texas at Austin, Austin, TX 78712.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. The work of the first author was supported by the National Science Foundation under Grant No. MCS-7927062. The work of the second author was supported in part by the Office of Naval Research under Contract N00014-79-C-0808. The third author was supported in part by the Joint Services Electronics Program under Contract F49620-77-C-0101.

SIGNIFICANCE AND EXPLANATION

A common problem in the analysis of stochastic systems is the estimation of the system's state given only noise-corrupted or incomplete observations. For instance, examples occur in communications theory when one wants to estimate a signal sent over a noisy channel. The problem of filtering is to build an estimate, i.e. filter, that provides the best information about the state given the observations.

The most desirable solution to a filtering problem is a recursive, physically realizable algorithm that computes the best mean-square error estimate, and thus it is important to find models for which such algorithms exist. Recently, Beneš defined a class of filtering problems that allow explicit computation of the conditional density of the signal given the past of the observations. This paper extends his result and then uses it to build exact, recursive algorithms for estimating any moment of the signal and for estimating polynomial-type transformations of the signal.

Accession TAB		<input checked="checked" type="checkbox"/>
GEN&I		<input type="checkbox"/>
TAB		<input type="checkbox"/>
Announced		<input type="checkbox"/>
Classification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or	Special
A		

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

EXPLICIT FILTERS FOR DIFFUSIONS WITH CERTAIN NONLINEAR DRIFTS

D. L. Ocone¹, J. S. Baras² and S. I. Marcus³

1. Introduction.

Let $f(x)$ be a real-valued function defined on all of \mathbb{R} and satisfying the Riccati equation

$$f'(x) + f^2(x) = ax^2 + bx + c. \quad (1.1)$$

It is assumed that f has no singularities. Note that this implies $a > 0$, for otherwise f explodes at some finite x . This paper considers the filtering problem

$$dx(t) = f(x(t))dt + db(t) \quad (1.2)$$

$$x(0) = x \in \mathbb{R}$$

$$dy(t) = x(t)dt + dw(t) \quad (1.3)$$

in which $b(\cdot)$ and $w(\cdot)$ are independent Wiener processes, $x(\cdot)$ is the signal, and $y(\cdot)$ the observation of $x(\cdot)$.

For the system (1.1)-(1.3), Beneš [1] recently derived an explicit formula for the conditional density of $x(t)$ given F_t^y , where $F_t^y := \sigma$ -algebra generated by $\{y(s) | 0 \leq s \leq t\}$. This result is interesting because the class of functions satisfying (1.1) includes nonlinear f , whereas conditional densities for (1.2)-(1.3) had been computed previously only for

¹Mathematics Department and the Mathematics Research Center, University of Wisconsin-Madison, Madison, WI 53706.

²Electrical Engineering Department, University of Maryland, College Park, MD 20742.

³Department of Electrical Engineering, University of Texas at Austin, Austin, TX 78712.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. The work of the first author was supported by the National Science Foundation under Grant No. MCS-7927062. The work of the second author was supported in part by the Office of Naval Research under Contract N00014-79-C-0808. The third author was supported in part by the Joint Services Electronics Program under Contract F49620-77-C-0101.

linear f : for examples and an extension to the multidimensional case, see Beneš [1].

Besides conditional densities one also wants to calculate filters $E\{\varphi|F_t^Y\}$ of $x(\cdot)$ -dependent statistics φ . Given a random process $\varphi(t)$ we shall say that $E\{\varphi(t)|F_t^Y\}$ is finite dimensionally computable (FDC) if it can be expressed as the output of a finite dimensional system of stochastic differential equations driven by $y(\cdot)$. When the signal dynamics in (1.2) are linear, many examples of FDC estimates are known. For example, $E\{x^n(t)|F_t^Y\}$ is FDC for any integer n , since the Kalman-Bucy equations calculate

$E\{x(t)|F_t^Y\}$ and the conditional variance, and higher order moments are derived from these by virtue of the normality of the conditional density. A more subtle class of examples consists of estimates $E\{\eta(t)|F_t^Y\}$ where $\eta(t)$ is any polynomial functional of $x(\cdot)$ in the form

$$\eta(t) = \int_0^t \dots \int_0^{s_{n-1}} \gamma(s_1, \dots, s_n) x^{k_1}(s_1) \dots x^{k_n}(s_n) ds_n \dots ds_1, \quad 1 \leq i \leq n,$$

in which γ is a separable function and the $\{k_i\}$ are non-negative integers (Marcus, Willsky [7]; Marcus, Mitter, Ocone [6]). Formulae and recursive systems for the smoothed estimate $E\{x(s)|F_t^Y\}$ are also well known (see [3]).

In this note, we extend the linear theory by showing that these same statistics are FDC for the general model (1.1)-(1.3). The strategy, as in the linear case, is to derive finite dimensional systems by using the explicit form of the conditional density to truncate formally infinite dimensional systems of moment equations. The material is organized as follows. In §2 we calculate conditional joint densities of $x(\cdot)$ given F_t^Y . As a consequence, we show that the conditional law of the process $\{x(s)|0 \leq s \leq t\}$ given F_t^Y and $x(t)$ is Gaussian. This is precisely the feature that makes it possible to handle polynomial functionals. In §3, we prove FDC of conditional moments, smoothers, and polynomial functionals.

Lie algebraic techniques from geometric control theory have been introduced recently into filtering, especially as regards finite dimensional computability, and they have been worked out successfully for known FDC problems in which f is linear. (For a survey of these ideas, see Brockett [8].) The main results of this paper, in particular proposition 3.7, were suggested by a Lie algebraic analysis of (1.1)-(1.3). Consequently, it seems that the full range of the Lie theory for linear drifts extends to the general case (1.1). Since our methods here are not algebraic, we do not pursue the issue further, but refer instead to Ocone [7] for further discussion.

2. Conditional Joint Densities.

Let $x(\cdot)$ and $y(\cdot)$ be given from (1.1)-(1.3), and let $t = s_0 > s_1 > \dots > s_n > 0$, $\underline{z} = (z_0, z_1, \dots, z_n)^T$. The expression $p(z_0, t; z_1, s_1, \dots, z_n, s_n | F_t^Y)$ shall denote the joint density of $(x(t), x(s_1), \dots, x(s_n))$ conditioned on F_t^Y , that is, for any bounded, Borel $\psi : R^{n+1} \rightarrow R$

$$E\{\psi(x(t), \dots, x(s_n)) | F_t^Y\} = \int_{R^{n+1}} \psi(\underline{z}) p(z_0, t; z_1, s_1, \dots, z_n, s_n | F_t^Y) d\underline{z}.$$

In Theorem 2.1 of this section, we employ a method of Beneš [1] to compute an explicit formula for this conditional density. From this we then derive

$p(z_1, s_1, \dots, z_n, s_n | F_t^Y, x(t))$ the conditional density of $(x(s_1), \dots, x(s_n))$ given F_t^Y and $x(t)$.

The results are stated in terms of an auxiliary process $\xi(t)$, evolving in R^3 and defined by

$$d\xi(t) = A(t)\xi(t)dt + \begin{bmatrix} 1 \\ y(t) \\ 0 \end{bmatrix} dB(t)$$

$$\xi(0) = (x, 0, 0)^T$$

where

$$A(t) = \begin{bmatrix} -\kappa & 0 & 0 \\ 0 & 0 & 0 \\ \kappa y(t) - \frac{1}{2}b & 0 & 0 \end{bmatrix},$$

$$\kappa = (a+1)^{1/2}$$

and $B(\cdot)$ is a Brownian motion independent of the signal and observation noises, $b(\cdot)$ and $w(\cdot)$.

Let

$$\Xi_0 = (\xi^T(t), \xi^T(s_1), \dots, \xi^T(s_n))^T$$

$$\Xi = (\xi_1(t), \xi_1(s_1), \dots, \xi_1(s_n))^T$$

Then the following conditional moments are needed:

$$m(t) := E\{\xi(t) | F_t^Y\}$$

$$R(t, s) := \text{cov}(\xi(t), \xi(s) | F_t^Y)$$

$$R(t) = [r_{ij}(t)]_{1 \leq i, j \leq 3} := R(t, t)$$

$$M(t, s_1, \dots, s_n) := E\{\Xi | F_t^Y\}$$

$$= (m_1(t), \dots, m_1(s_n))^T$$

$$P_0(t, s_1, \dots, s_n) := \text{Var}(\Xi_0 | F_t^Y)$$

$$P(t, s_1, \dots, s_n) := \text{Cov}(\Xi, \Xi_0 | F_t^Y)$$

$$= E\{(\Xi - M)(\Xi_0 - E\{\Xi_0 | F_t^Y\})^T | F_t^Y\}$$

$$Q(t, s_1, \dots, s_n) := \text{Var}(\Xi | F_t^Y).$$

To simplify later expressions, we shall often drop the (t, s_1, \dots, s_n)

dependence and write only M, P_0, P and Q . In addition, let

$v = (0, 1, -1, 0, \dots, 0)^T \in \mathbb{R}^{3(n+1)}$. The random vectors $P_0 v$ and Pv play an important role in theorem 2.1 and are related by

$$Pv = ((P_0 v)_1, (P_0 v)_4, \dots, (P_0 v)_{3n+1})^T.$$

Note that $(Pv)_k = (P_0 v)_{3k-2} = \text{cov}(\xi_1(s_{k-1}), \xi_2(t) - \xi_3(t) | F_t^Y)$.

It is important to observe that all these conditional moments are properly thought of as functionals (on $C[0, t]$) of $y(\cdot)$. Indeed, these functionals are easily calculated by solving for every $\bar{y}(\cdot) \in C[0, t]$ the system

$$d\bar{\xi}_{\bar{y}}(s) = A_{\bar{y}}(s) \bar{\xi}_{\bar{y}}(s) ds + \begin{bmatrix} 1 \\ \bar{y}(s) \\ 0 \end{bmatrix} dB(s)$$

$$\bar{\xi}_{\bar{y}}(0) = (x, 0, 0)^T.$$

Then, if $\bar{m}(t, \bar{y}(\cdot)) : E \bar{\xi}_{\bar{y}}(t)$, $m(t)(\omega) = \bar{m}(t, y(\cdot)(\omega))$, and similiary for R , P_0 , etc.

As a final bit of notation, set $F(z_0) = \int_0^{z_0} f(s) ds$.

Theorem 2.1. Let $t > s_1 > s_2 > \dots > s_n$. Then

$$p(z_0, t; z_1, s_1; \dots; z_n, s_n | F_t^y) \\ = (1/\Psi) \exp \{ F(z_0) + z_0 y(t) + x \kappa(z_0 - x) + \kappa \left[\frac{(z_0 - x)^2}{2} - t \right] \}$$

$$\times \exp \{ -1/2 \langle \underline{z} - M + P v, Q^{-1}(\underline{z} - M + P v) \rangle \}$$

where $\Psi = \Psi(t, s_1, \dots, s_n, x)$ is a normalizing factor.

Proof. The demonstration is analogous to Beneš' proof in [1] of the case

$p(z_0, t | F_t^y)$, and so we shall only sketch the main steps. The Kallianpur-Striebel formula for conditional estimation in system (1.2)-(1.3) implies that

$$p(z_0, t; \dots; z_n, s_n | F_t^y) dz_0 \dots dz_n \\ \propto E \left\{ \prod_{j=0}^n 1_{\{x(s_j) \in dz_j\}} \exp \int_0^t x(s) dy(s) - 1/2 \int_0^t x^2(s) ds \right\} \quad (2.2)$$

In expression (2.2), we think of $y(t)$ as a fixed function and of \tilde{E} as an expectation against the measure μ induced by (1.2) on the space of continuous sample paths $x(\cdot)$. ' \propto ' here means proportional up to a normalization factor that does not depend on \underline{z} . To evaluate (2.2) we follow Beneš and apply a sequence of Girsanov measure transformations. Let $\mu_{B,x}$ denote the measure on $C[0, t]$ induced by $x + B(\cdot)$. Then $\mu \ll \mu_{B,x}$ and

$$\frac{d\mu}{d\mu_{B,x}}(x(\cdot)) = \exp\left\{\int_0^t f(x(s))dx(s) - \frac{1}{2}\int_0^t f^2(x(s))ds\right\}.$$

By using this to change measures in the expectation term of (2.2) and noting, by Ito's rule, that

$$\int_0^t x(s)dy(s) = x(t)y(t) - \int_0^t y(s)dx(s)$$

and

$$F(x+B(t)) - F(x) = \int_0^t f(x+B(s))dB(s) - \frac{1}{2}\int_0^t f'(x+B(s))ds,$$

we derive

$$p(\cdot\cdot\cdot)dz = \exp\{F(z_0) - F(x) + z_0 y_t\} \quad (2.6)$$

$$\times E\left\{\prod_{j=0}^n 1_{\{x+B(s_j)\}} \varepsilon dz_j\right\} \exp\left\{-\int_0^t y(s)dB(s) - \frac{1}{2}\int_0^t V(x+B(s))ds\right\}$$

where $V(x) = (a+1)x^2 + bx + c$. To evaluate this last expectation, we treat the quadratic part of V as arising from the Radon-Nikodym derivative of $x + B(\cdot)$ with respect to the Ornstein-Uhlenbeck process $\xi_1(\cdot)$; the linear terms in the exponent in (2.6) can then be re-expressed in terms of $\xi_2(\cdot)$ and $\xi_3(\cdot)$. The result is

$$\exp\left\{xk(z_0-x) + k\left[\frac{(z_0-x)^2}{2} - t\right]\right\} \quad (2.7)$$

$$\times E\left\{\prod_{j=0}^n 1_{\{\xi_1(s_j)\}} \varepsilon dz_j\right\} \exp\left\{-(\xi_2(t) - \xi_3(t))\right\} F_t^y.$$

Given F_t^y , $\xi(\cdot)$ is a Gaussian process, and thus the expectation in (2.7) may be written, up to a normalizing factor, as

$$\int_{R^{2n+2}} d\bar{\zeta} \exp\left\{-\frac{1}{2}\langle \zeta - E_y \bar{\zeta}_0 + P_0 v, P_0^{-1}(\zeta - E_y \bar{\zeta}_0 + P_0 v) \rangle\right\} \exp\frac{1}{2}\langle v, P_0 v \rangle$$

where $E_y \Xi_0 = E\{\Xi_0 | F_t^Y\}$ and where $d\bar{\zeta}$ signifies that $\zeta_1 = z_0$,

$\zeta_4 = z_1, \dots, \zeta_{3n+1} = z_n$ are held fixed and integration is over the remaining variables. But this last expression integrates by standard Gaussian integral formulae to a factor proportional to

$$\exp\{-\frac{1}{2} \langle \underline{z} - M + Pv, Q^{-1}(\underline{z} - M + P) \rangle\}. \quad (2.8)$$

By combining (2.6)-(2.8), we arrive at the desired result. ■

The conditional density of the process $x(\cdot)$ thus consists of a Gaussian factor multiplied by $\exp F(z_0)$. Further conditioning on $x(t)$ will remove $\exp F(z_0)$ and leave only the normal part. Indeed, let

$$\Xi^{(2)} = (\xi_1(s_1), \dots, \xi_1(s_n))^T$$

$$Q(t, s_1, \dots, s_n) = \begin{bmatrix} r_{11}(t) & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$$

$$Q_{22}(t, s_1, \dots, s_n) = \text{Var}(\Xi^{(2)} | F_t^Y)$$

$$Pv = \begin{bmatrix} r_{12}(t) - r_{13}(t) \\ P_v^{(2)} \end{bmatrix}, \quad M = \begin{bmatrix} m_1(t) \\ M^{(2)} \end{bmatrix}$$

Corollary 2.9. The conditional law of $(x(s_1), \dots, x(s_n))$ given F_t^Y and $x(t)$ is normal with mean

$$M^{(2)} - Pv^{(2)} + r_{11}^{-1}(t) Q_{21}(x(t) - m_1(t) + Pv_1)$$

and variance $Q_{22} - r_{11}^{-1}(t) Q_{21} Q_{12}$.

Proof. Use (2.1) to conclude that

$$p(z_1, s_1; \dots; z_n, s_n | x_0(t) = z_0, F_t^y) = p(z_0, t; \dots; z_n, s_n | F_t^y) / p(z_0, t | F_t^y)$$

$$\propto \exp\{-\frac{1}{2} \langle \underline{z} - M + Pv, Q^{-1}(\underline{z} - M + Pv) \rangle\}$$

$$\times \exp\{\frac{1}{2} \frac{1}{r_{11}(t)} (z_0 - m_1(t) + (Pv)_1)^2\}.$$

But this is just the conditional density of η_1, \dots, η_n given η_0 where

(η_0, \dots, η_n) is a normal random vector with mean $M - Pv$ and variance Q .

The result then follows from the standard formula for conditioning one part of a normal random vector upon another. ■

This corollary demonstrates how closely those diffusions defined by (1.1)-(1.2) are related to Gaussian processes. In fact, the steps above can be repeated to calculate joint densities of the process $x(\cdot)$; (2.1) yields the correct expression if $y(\cdot)$ is replaced by 0. Then, in the same way, it follows that $(x(s_1), \dots, x(s_n) | x(t))$ is normal, that is, the process $\{x(s) | 0 \leq s \leq t\}$ conditioned on the endpoint $x(t)$ is Gaussian. This conditional normality is key to the filtering results of §3.

3. Filtering Equations.

3.1. Conditional Moments.

Let $\hat{x}^n(t) := E\{x(t) | F_t^Y\}$. $\hat{x}^n(t)$ satisfies the equation

$$d\hat{x}^n = [n(n-1)/2 \hat{x}^{n-2} + nf(x)\hat{x}^{n-1}]dt + [\hat{x}^{n+1} - \hat{x}\hat{x}^n][dy - \hat{x}dt] \quad (3.1)$$

(Fujisaki, Kallianpur, Kunita [2]). To calculate \hat{x}^n one must therefore also find \hat{x}^{n+1} , $f(x)\hat{x}^n$, etc. These in turn also satisfy stochastic differential equations that introduce yet other quantities to be estimated. Continuing in this manner, if we begin with \hat{x} , we arrive at an infinite, coupled set of conditional moment equations for $\hat{x}, \hat{x}^2, \dots, \hat{x}^n, \dots, f(x)\hat{x}$, etc. Our approach to finite dimensional computability will be to use conditional moment identities, derived from the form of $p(x,t | F_t^Y)$, to truncate this infinite system after a finite number of terms. Actually, FDC of the moments could be argued on general grounds using the fact that the FDC process $\mu(t)$ (see immediately below) characterizes $p(z,t | F_t^Y)$. However, the approach here leads to explicit filtering equations.

Straightforward analysis of (2.1) shows that

$$p(x,t | F_t^Y) = \frac{1}{N(t,x)} \exp\{F(z) - (z - \mu(t))^2 / 2\sigma(t)\}$$

where $N(t,x)$ is a normalizing factor, and

$$\dot{\sigma} = 1 - \kappa^2 \sigma^2 \quad \sigma(0) = 0 \quad (3.2)$$

$$d\mu = [-\kappa^2 \sigma \mu - 1/2 b \sigma]dt + \sigma dy \quad \mu(0) = x \quad (3.3)$$

(See Beneš [1].) Note that (3.2) implies that $1 - \kappa^2 \sigma^2(t) > 0 \quad \forall t > 0$,

and hence that

$$a < \sigma^{-1}(t) \quad \forall t > 0.$$

As a consequence, $\widehat{x^n}(t)$ is well-defined for any n . If $a > 0$, (1.1)

implies $f(x) \sim a^{1/2} x$ as $|x| \rightarrow \infty$ and thus that $F(z) \sim \frac{a^{1/2}}{2} z^2$ as $|z| \rightarrow \infty$. Therefore

$$F(z) - (z - \mu)^2/2\sigma \sim 1/2(a^{1/2} - \sigma^{-1})z^2, \quad |z| \rightarrow \infty$$

and so $P(z, t | F_t^y)$ decays as $\exp[-\delta z^2]$, $\delta = 1/2(\sigma^{-1} - a^{1/2}) > 0$. If $a = 0$, then $b = 0$, $c > 0$ is necessary in order that (1.1) have a solution f on \mathbb{R} without singularities. Then f will be bounded, and $F(z)$ will grow at most linearly.

Lemma 3.4. For $n > 0$

$$\begin{aligned} (\sigma^{-1} - a)\widehat{x^{n+2}} &= (b + 2\mu\sigma^{-2})\widehat{x^{n+1}} + (c + ((2n+1) - \mu^2)\sigma^{-1})\widehat{x^n} \\ &\quad + 2\mu\sigma^{-1}\widehat{x^{n-1}} - n(n-1)\widehat{x^{n-2}}. \end{aligned}$$

Proof. Integrating by parts, we have

$$\begin{aligned} \frac{1}{N(t, x)} \int_{-\infty}^{\infty} dz \exp F(z) \frac{d^2}{dz^2} (z^n \exp[-(z - \mu)^2/2\sigma]) \\ = \int_{-\infty}^{\infty} dz [f'(z) + f^2(z)] z^n p(z, t | F_t^y) \\ = a\widehat{x^{n+2}} + b\widehat{x^{n+1}} + c\widehat{x^n}. \end{aligned}$$

Evaluation of $\frac{d^2}{dz^2} (z^n \exp[-(z - \mu)^2/2\sigma])$ in the first integral leads to the desired identity.

Proposition 3.5: $\hat{x}^n(t)$ is FDC for all n .

Proof: Since $\sigma^{-1}(t) - a > 0$, lemma 3.1 implies that $\hat{x}^m(t)$, $m > 2$ can be expressed as a linear combination with FDC coefficients of lower order conditional moments. Thus it suffices to prove that $\hat{x}(t)$ is FDC. For $\hat{x}(t)$, (3.1) becomes

$$d\hat{x} = \widehat{f(x)}dt + (\hat{x}^2 - \hat{x}^2)(dy - \hat{x}dt), \quad \hat{x}(0) = x.$$

Now

$$\begin{aligned} \widehat{f(x)} &= \frac{1}{N(t,x)} \int dz [d/dz \exp F(z)] \exp[-(z - \mu)^2/2\sigma] \\ &= - \int dz (z - \mu) \sigma^{-1} p(z, t | F_t^y) \\ &= -(\hat{x} - \mu) \sigma^{-1} \end{aligned}$$

and

$$\hat{x}^2 = [\hat{x}(b + 2\mu\sigma^{-2}) + c + (1 - \mu^2)\sigma^{-1}](\sigma^{-1} - a)^{-1}.$$

Thus

$$\begin{aligned} d\hat{x} &= (\mu - \hat{x}) \sigma^{-1} dt \\ &+ [(\hat{x}(b + 2\mu\sigma^{-2}) + c + (1 - \mu^2)\sigma^{-1})(\sigma^{-1} - a)^{-1} - \hat{x}^2][dy - \hat{x}dt] \end{aligned}$$

together with (3.2)-(3.3) constitutes a finite-dimensional system for $\hat{x}(t)$.

Remark. It is also clear that $\widehat{f(x(t))x^n(t)}$ will be FDC for any n ; simply eliminate $f(x(t))$ in favor of polynomials in $x(t)$ by integration by parts as in the proof above. In the same way, filters may be constructed for any conditional estimate in the infinite set of moment equations generated by starting with $\hat{x}(t)$ and using (3.1).

3.2. Smoothing.

Proposition 3.6. Let $s < t$. Then

$$E\{x(s)|F_t^Y\} = \frac{\sinh ks}{\sinh kt} [\hat{x}(t) - xe^{-kt} + (P(t,s)v)_1] + xe^{-ks} - (P(t,s)v)_2.$$

Proof: This is immediately consequent from corollary 2.9 once it is noted that

$$r_{11}^{-1}(t)Q_{21} = r_{11}^{-1}(t)\text{cov}(\xi_1(t), \xi_1(s)) = \frac{\sinh ks}{\sinh kt}$$

and

$$m_1(t) = xe^{-kt}.$$

Explicit formulae for $(P(t,s)v)_1$ and $(P(t,s)v)_2$ are easily found and will be given in the next section.

3.3. Polynomial Functionals.

Let $\eta(t)$ be any non-anticipating functional of the signal process of the form

$$\eta(t) = \int_0^t \dots \int_0^{s_{n-1}} \gamma(s_1, \dots, s_n) x^{k_1}(s_1) \dots x^{k_n}(s_n) ds_n \dots ds_1$$

where k_1, \dots, k_n are arbitrary, non-negative integers and $\gamma(s_1, \dots, s_n)$ is a separable function. Also let $\hat{\eta}(t) = E\{\eta(t)|F_t^Y\}$.

Proposition 3.7. $\hat{\eta}(t)$ is FDC

For simplicity we restrict attention to the case

$k_1 = k_2 = \dots = k_n = 1$; the method of proof extends easily to general choices for the k_i 's. Our proof relies on the following identity, presented here in the form that it appears in Marcus and Willsky [7].

Lemma 3.8. Let (u_1, \dots, u_m) be a normal random vector with $e_j = Eu_j$ and $V_{ij} = \text{cov}(u_i, u_j)$. Then

$$E[u_1 \dots u_m] = \prod_1^m e_j + \sum V_{j_1 j_2} e_{j_3} \dots e_{j_m} + \sum V_{j_1 j_2} V_{j_3 j_4} e_{j_5} \dots e_{j_m} + \dots$$

The sums are taken over all possible combinations of pairs of indices.

Lemma 3.9. Let $s_1 > \dots > s_n$. Then

$$E\{x(s_1) \dots x(s_n) | F_t^Y, x(t)\} = \sum_{j=0}^n x^j(t) a_j(t, s_1, \dots, s_n)$$

for some separable functions $a_j(t, s_1, \dots, s_n)$, $0 \leq j \leq n$, depending on $y(\cdot)$.

Proof. Let $\ell_j = M_j^{(2)} - P v_j^{(2)} + r_{11}^{-1}(t) Q_{12}(t, \dots, s_n) [x(t) - m_1(t) + (Pv)_1]$

$$\tilde{Q} = Q_{22} - r_{11}^{-1}(t) Q_{12} Q_{21}$$

and apply lemma 3.8 and corollary 2.9. Thus

$$E\{x(s_1) \dots x(s_n) | x(t), F_t^Y\} = \ell_1 \dots \ell_n + \sum \tilde{Q}_{j_1 j_2} \ell_{j_3} \dots \ell_{j_n} \\ + \sum \sum \tilde{Q}_{j_1 j_2} \tilde{Q}_{j_3 j_4} \ell_{j_5} \dots \ell_{j_n} + \dots$$

This is an n^{th} order polynomial in $x(t)$ since ℓ_j is a linear function of $x(t)$ for each j . Moreover, it is clear that the coefficients a_j will be separable if $P_0(t, s_1, \dots, s_n)$ and hence $Q(t, s_1, \dots, s_n)$ are separable (Q is a submatrix of P_0). But

$$P_0(t, s_1, \dots, s_n) = \begin{bmatrix} R(t, t) & R(t, s_1) & \dots & R(t, s_n) \\ R(s_1, t) & R(s_1, s_1) & \dots & R(s_1, s_n) \\ \vdots & & & \\ R(s_n, t) & \dots & \dots & R(s_n, s_n) \end{bmatrix}$$

and $R(t, s) = \phi_A(t, s)R(s, s) = \phi_A(t, 0)\phi_A^{-1}(s, 0)R(s)$ where $\phi_A(t, s)$ is the state transition matrix of $A(t)$. Thus $P_0(t, \dots, s_n)$ is indeed separable.

Remark. Upon further inspection of the terms $\phi_A(t, s)$ and $R(s)$, the proof of this lemma demonstrates that a_j may be written

$$a_j(t, s_1, \dots, s_n) = \sum_1 \beta_{j,1}^1(t) \beta_{j,2}^1(s_1) \dots \beta_{j,n}^1(s_n) \quad (3.10)$$

where each $\beta_{j,k}^1$, $0 < k < n$ is either deterministic or a non-anticipating functional of the observation process $y(\cdot)$.

Proof of proposition: By using lemma 3.9

$$E\{x(t) | F_t^Y\}$$

$$\begin{aligned} &= \int_0^t \dots \int_0^{s_{n-1}} \gamma_1(s_1) \dots \gamma_n(s_n) E\{E[x(s_1) \dots x(s_n) | F_t^Y, x(t)] | F_t^Y\} ds_n \dots ds_1 \\ &= \sum_{j=0}^n \hat{x}^j(t) \int_0^t \int_0^{s_{n-1}} \gamma(s_1, \dots, s_n) a_j(t, s_1, \dots, s_n) ds_n \dots ds_1. \end{aligned}$$

To complete the proof, it is only necessary to show that the coefficients of $\hat{x}^j(t)$ are FDC. Each coefficient is a sum of terms of the form

$$u_{n+1}(t) = \alpha_0(t) \int_0^t \dots \int_0^{s_n} \alpha_1(s_1) \dots \alpha_n(s_n) ds_n \dots ds_1$$

which can be computed on-line by the system

$$\begin{aligned}
\dot{u}_1(t) &= \alpha_n(t) & u_1(0) &= 0 \\
\dot{u}_2(t) &= \alpha_{n-1}(t)u_1(t) & u_2(0) &= 0 \\
&\vdots & & \\
\dot{u}_n(t) &= \alpha_1(t)u_n(t) & u_n(0) &= 0 \\
\dot{u}_{n+1}(t) &= \alpha_0(t)u_n(t) & u_{n+1}(0) &= 0.
\end{aligned}$$

Thus $u_{n+1}(t)$ will be FDC if each $\alpha_i(t)$ is FDC. However, reasoning from the remark after lemma 3.9, each $\alpha_i(t)$ will be either deterministic, or a deterministic function multiplied by one of the $y(\cdot)$ -dependent $\beta_{j,k}^i(t)$ from (3.10). Now by the proof of lemma 3.9 these $\beta_{j,k}^i(t)$ come from the $y(\cdot)$ -dependent terms in ℓ_i , $1 \leq i \leq n$, and \tilde{Q} . Actually, inspection reveals that ¹⁾ \tilde{Q} is independent of $y(\cdot)$, since it is a function of the joint covariance $\xi_1(t), \dots, \xi_1(s_n)$; and ²⁾ the only y -dependence in the ℓ_i lies in P_v . Recall from §2, that a typical element of P_v is $\text{cov}(\xi_1(s), \xi_2(t) - \xi_3(t))$; a simple calculation shows this equals

$$\begin{aligned}
\text{cov}(\xi_1(s), \xi_2(s) - \xi_3(s)) &= \kappa^{-1} \sinh \kappa s \int_0^t e^{-\kappa u} g(u) du \\
&= \kappa^{-1} \sinh \kappa s \int_0^s e^{-\kappa u} g(u) du
\end{aligned} \tag{3.11}$$

where $g(u) = \kappa y(u) - \frac{1}{2}b$. The $y(\cdot)$ functionals in this expression are certainly FDC and thus the components $\beta_{j,k}^i(t)$ are FDC. This completes the proof.

The proof of proposition 3.7 is similar to the proof of the linear case due to Marcus and Willsky [7] in its use of Gaussian moment identities. Actually, because conditional Gaussianity obtains in the linear case without first conditioning on $x(t)$, Marcus and Willsky are able to use the general filtering equation and a simpler moment identity than lemma 3.8 to build a

proof by induction on the order of $\eta(t)$. By 'general filtering equation' is meant the representation of Fujisaki, Kallianpur, Kunita [2]:

$$\begin{aligned} d\hat{\eta} = & E\{x^{k_1}(t) \int_0^t \dots \int_0^{s_{n-2}} \gamma(t, \dots) x^{k_2}(s_1) \dots x^{k_n}(s_{n-1}) ds_{n-1} \dots ds_1 | F_t^Y\} dt \\ & + [\widehat{\eta(t)x(t)} - \hat{\eta}(t)\hat{x}(t)][dy(t) - \hat{x}(t)dt]. \end{aligned} \quad (3.12)$$

In our proof, application of (3.12) is superfluous, although, when calculating the filter in a particular example, it can be employed to advantage. The example below will illustrate the possibilities.

Marcus, et al [6] give an alternate proof of finite dimensional computability when f is linear by using homogeneous chaos theory and multiple integral expansions. Such an approach might also be possible here by first conditioning on $x(t)$, but this is not pursued, since the calculation would ultimately be like the one here.

Example. Consider the model

$$\begin{aligned} dx(t) &= f(x(t))dt + db(t) & x(0) &= x_0 \\ d\eta(t) &= x^2(t)dt & \eta(0) &= 0 \\ dy(t) &= x(t)dt + dw(t) & y(0) &= 0 \end{aligned}$$

where the $x(\cdot)$ and $y(\cdot)$ equations are as in (1.1)-(1.2). We will present a finite dimensional system for computing $\hat{\eta}(t) = E\{\int_0^t x^2(s)ds | F_t^Y\}$. This problem was chosen in part by way of comparison to the special case $f \equiv 0$ which is treated in detail in Liu and Marcus [4]. The system given here for $\hat{\eta}(t)$ is, of course, one among many possibilities; our construction was guided by the decision to use Ito equations driven by the innovations $dv(t) := dy(t) - \hat{x}(t)dt$.

Let

$$u(t) = (\sinh \kappa t)^{-2} \int_0^t (\sinh \kappa s)^2 ds = (\sinh \kappa t)^{-1} [(4\kappa)^{-1} \sinh 2\kappa t - \frac{1}{2} t].$$

Then

$$\begin{aligned} d\hat{n} &= \hat{x}^2 dt + u(t) [\hat{x}^3 - \hat{x}^2 \hat{x}] dv \\ &+ 2[\hat{x}^2 - \hat{x}^2] [u(t)(v_2(t)\kappa^{-1} \sinh \kappa t + \frac{\sigma}{1+\kappa\sigma} (\sigma y(t) - \mu)) \\ &\quad + v_1(t) \sinh \kappa t - (\kappa \sinh \kappa t)^{-1} v_3(t)] dv \\ \hat{n}(0) &= 0 \\ dv_1 &= \sinh \kappa t (\mu - \sigma y(t)) \frac{\sigma}{1+\kappa\sigma} dt & v_1(0) &= 0 \\ dv_2 &= e^{-\kappa t} [\kappa y(t) - \frac{1}{2} b] dt & v_2(0) &= 0 \\ dv_3 &= (\sinh \kappa t)^2 v_2(t) dt & v_3(0) &= 0. \end{aligned} \quad (3.13)$$

(3.13) is not the complete system, since equations for $\mu, \sigma, \hat{x}, \hat{x}^2$, and \hat{x}^3 are also needed. However these are easily garnered from §3. A brief derivation of (3.13) follows. From (3.12)

$$d\hat{n} = \hat{x}^2 dt + (\hat{n}\hat{x} - \hat{n}\hat{x}) dv. \quad (3.14)$$

However, corollary 2.9 implies

$$\begin{aligned} E\{\hat{x}^2(s) | F_{t,x(t)}^y\} &= Q_{22} - r_{11}^{-1}(t) Q_{21} Q_{12} + E^2\{x(s) | F_{t,x(t)}^y\} \\ &= F(t,s) + r_{11}^{-2}(t) Q_{21}^2 x^2(t) \\ &\quad + 2x(t) r_{11}^{-1}(t) Q_{21} [m_1(s) - (Pv)_2 - r_{11}^{-1}(t) Q_{21} (m_1(t) - Pv_1)] \end{aligned} \quad (3.15)$$

where $F(t,s)$ combines those terms not depending on $x(t)$. Now

$$r_{11}^{-1}(t)Q_{21} = \sinh \kappa s / \sinh \kappa t,$$

and

$$Pv_4 = Pv_1(s) - \kappa^{-1} \sinh \kappa s \left[\int_0^t e^{-\kappa u} g(u) du - \int_0^s e^{-\kappa u} g(u) du \right].$$

(Recall that $g(u) = \kappa y(u) - 1/2$ b.) Further, it can be shown

$$m_1(t) - Pv_1(t) = \frac{\sigma}{1+\kappa\sigma} (\mu - \sigma y)(t).$$

Using these identities in (3.15) it follows that

$$\begin{aligned} \hat{n}_x - \hat{\bar{n}}_x &= \int_0^t E\{(x(t) - \hat{x}(t))E(x^2(s)|F_t^y, x(t))|F_t^y\} ds \\ &= (\hat{x}^3 - \hat{x}^2 \hat{x})u(t) + 2(\hat{x}^2 - \hat{x}^2)u(t) \left[(\sigma y - \mu) \frac{\sigma}{1+\kappa\sigma} + \kappa^{-1} \sinh \kappa t \int_0^t e^{-\kappa u} g(u) du \right. \\ &\quad \left. + 2(\hat{x}^2 - \hat{x}^2) [(\sinh \kappa t)^{-1} \int_0^t (\sinh \kappa s) [m_1(s) - Pv_1(s)] ds \right. \\ &\quad \left. - (\kappa \sinh \kappa t)^{-1} \int_0^t (\sinh \kappa s)^2 \int_0^s e^{-\kappa u} g(u) du ds \right]. \\ &= (\hat{x}^3 - \hat{x}^2 \hat{x})u(t) + 2(\hat{x}^2 - \hat{x}^2) [u(t) (v_2(t) \kappa^{-1} \sinh \kappa t + \frac{\sigma}{1+\kappa\sigma} (\sigma y(t) - \mu)) \\ &\quad + v_1(t) \sinh \kappa + -(\kappa \sinh \kappa t)^{-1} v_3(t)] \end{aligned}$$

(3.16)

Placing (3.16) in (3.14) one obtains the desired result.

REFERENCES

1. V. E. Beneš, Exact finite dimensional filters for certain diffusions with nonlinear drift, Stochastics 5 (1981), 65-92.
2. M. Fujisaki, G. Kallianpur, and H. Kunita, Stochastic differential equations for the non linear filtering problem, Osaka J. Math. 9 (1972), 19-40.
3. A. H. Jazewinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.
4. C-H Liu and S. I. Marcus, The Lie algebraic structure of a class of finite dimensional nonlinear filters, in Lectures in Applied Mathematics 18 (1980), 277-297.
5. S. I. Marcus, S. K. Mitter, and D. Ocone, Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time, in Analysis and Optimisation of Stochastic Systems (Proc. Internat. Conf. Univ. Oxford, Oxford, 1978), 387-406, Academic Press, London, 1980.
6. S. I. Marcus and A. S. Willsky, Algebraic structure and finite dimensional nonlinear estimation, SIAM J. Math. Anal. 9 (1978), 312-327.
7. D. Ocone, Finite dimensionally computable statistics and estimation algebras in nonlinear filtering, Proceedings, 1981 Int. Sym. Mathematical Theory of Networks and Systems, Santa Monica, CA, Aug. 5-7, 1981.
8. R. W. Brockett, to appear in Proceedings, NATO ASI on Mathematics of Stochastic Systems, Les Arcs, 1981, Reidel Press.

DLO:JSB:SIM/db

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2254	2. GOVT ACCESSION NO. AD-A103867	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EXPLICIT FILTERS FOR DIFFUSIONS WITH CERTAIN NONLINEAR DRIFTS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
7. AUTHOR(s) D. L. Ocone, J. S. Baras and S. I. Marcus		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		8. CONTRACT OR GRANT NUMBER(s) F49620-77-C-0101 N00014-79-C-0808 DAAG29-80-C-0041 MCS-7927062
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 4-Probability and Statistics
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE August 1981
		13. NUMBER OF PAGES 20
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES U. S. Army Research Office National Science Foundation P. O. Box 12211 Washington, D. C. 20550 Research Triangle Park North Carolina 27709		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Nonlinear filtering		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Let $x(t)$ be a diffusion satisfying the stochastic differential equation $dx(t) = f(x(t))dt + db(t)$, where $f'(x) + f^2(x) = ax^2 + bx + c$, $a > 0$. V. Beneš gave an explicit formula for the conditional density of $x(t)$ given $y(s)$, $0 \leq s \leq t$, where $y(s) = \int_0^s x(s) ds + w(t)$, when $w(\cdot)$ is		

20. Abstract (continued)

a Brownian process independent of $x(\cdot)$. This result is extended and then applied to derive recursive filtering equations for estimating conditional moments $E\{x^n(t)|y(s), 0 < s < t\}$, for estimating polynomial functionals of $x(\cdot)$, and for smoothing.

**DA
FILM**